LENTIVIRAL VECTORS CARRYING SYNTHETIC BI-DIRECTIONAL PROMOTERS AND USES THEREOF

The present invention relates to bidirectional promoters allowing efficient and coordinate expression of two or more genes, to gene transfer vectors containing these promoters, to particles transducing said vectors into a cell, to the use of said vectors for the delivery and expression of multiple genes in target cells, also for gene therapy, and for the manufacturing of medicaments.

TECHNICAL BACKGROUND

Expression of multiple transgenes within the same target cells is required for several gene transfer and therapy applications[1]. Gene-function studies are best performed by expressing cDNAs together with a marker gene; by this approach, genetically-modified cells can be identified and monitored in vitro and in vivo. Similarly, gene therapy applications can be improved by purification of gene-corrected cells before in vivo administration, taking advantage of coordinate expression of selectable markers. Genetically-modified cells can be amplified ex vivo or in vivo by introducing growth-promoting or drug-resistance genes together with the therapeutic gene, as recently shown by MGMT-mediated selection of transduced Hematopoietic Stem Cells (HSC) [2]; using this approach, the efficacy of gene therapy can be increased, and its application potentially extended to a wide spectrum of diseases[3, 4]. Conversely, genetically-modified cells expressing conditionally cytotoxic genes, together with the therapeutic gene, can be eliminated in vivo, if adverse events occur; this approach is used to control graft-versus-host disease following donor T-lymphocytes infusion to treat leukemia relapse[5]; it may also provide an important safety provision in HSC gene transfer, given the recent occurrence of leukemia related to vector integration in a successful clinical trial of X-linked Severe Combined ImmunoDeficiency[6]. Coordinate expression of more than one transgene is essential when the activity to be reconstituted by gene transfer depends on multiple subunits encoded by different genes, or requires the synergism of separate molecules. For instance, reconstitution of the dopamine biosynthetic pathway in striatal neurons of Parkinson's disease patients requires co-expression of tyrosine hydroxylase with

GTP-cyclohydrolase I and/or DOPA decarboxylase[7]; cancer gene therapy may require co-expression of multiple antigens and/or cytokines in antigen-presenting cells for immunotherapy, and of two T-cell receptor chains in T-cells engineered for adoptive transfer[8].

In spite of such well-recognized needs, reaching coordinate, high-level expression of multiple transgenes in the majority of target cells has been a significant challenge for gene transfer technology. Two different transgenes have been expressed by two separate vectors; yet, only a fraction of target cells was transduced by both vectors and a heterogeneous population of cells was obtained that expressed either one or two genes in different ratios, preventing reliable studies and/or efficacious applications. Alternatively, two or more transgenes have been expressed by different promoters within the same vector[9]; yet, different tissue specificity and mutual interference between promoters often prevented efficient co-expression in the same target cells[10]. Differential splicing generates multiple transcripts from the same promoter, but it has proven difficult to adapt to viral delivery of multiple transgenes[11]. Chimeric polyproteins that self-process co-translationally into separate components have been generated using the self-cleaving peptide of the Foot and Mouth Disease Virus 2A[12, 13]; however, application of this technology to multiple gene transfer has been limited until now because it requires sophisticated engineering, restricts both proteins to the same cellular compartment, and introduces sequence changes that may affect protein activity, stability, and immunogenicity.

The most satisfactory approach to multiple gene transfer until now has relied on using internal ribosome entry sites (IRES's)[14]. These sequences, identified in viral and cellular transcripts, control translation in a ^{mRNA}Cap-independent manner and, when inserted between two genes in a bicistronic messenger RNA, allow translation of the downstream gene. The authors tested the performance of different IRES's in the context of self-inactivating (SIN) lentiviral vectors (LVs), and found significant limitations of this approach.

WO 02/064804 describes bi-directional dual promoter complexes that are effective for enhancing transcriptional activity of transgenes in plants.

The bi-directional promoters of the invention include a modified enhancer region with at least two core promoters on either side of the modified enhancer in a divergent orientation. The application refers to gene expression in plants. In addition, the approach requires the duplication of tandem oriented enhancer sequences in a modified internal region of the construct, to be joined by two identical or homologous minimal promoters on either sides. The instant invention does not require duplication of enhancer or any other sequences in the efficient promoter of the bi-directional construct, nor are need that the core promoters on either sides of it to share at least 30% identity. Finally, tandem duplication may be incompatible with retro/lentiviral delivery.

US 6,388,170 discloses plant vectors, having bi-directional promoters, comprising a minimal promoter and a common promoter, wherein said minimal promoters is operably linked to said common promoter, in opposite orientation to said common promoter, and 5' to said common promoter. Promoter sequences derived from plants and plant-infecting viruses are disclosed dnd tested in plant cells or plant parts. Given the substantial evolutionary distance between plants and animals, US 6,388,170 does not teach how to engineer animal promoters for bidirectional activity and whether bi-directional promoters may effectively work in animal cells. In addition, US 6,388,170 does not teach how to engineer bi-directional promoters for gene expression in animals and in animal cells using the available gene transfer methods.

WO01/34825 discloses cell lines, plasmids and vectors useful for the production of recombinant viruses such as adenoviruses, which are useful in gene therapy. The cell lines, plasmids and vectors comprise inducible promoters, such as bi-directional promoters for the coordinate expression of bidirectionally cloned gene. However only bi-directional Tet-regulated constructs are disclosed.

Thus, the authors explored novel strategies to take full advantage of gene transfer systems, such as LV, that allow efficient ex vivo transduction and direct in vivo administration.

DESCRIPTION OF THE INVENTION

The authors developed a novel vector design in which synthetic bi-directional promoters mediated coordinate transcription of two divergent RNAs. The authors show that LVs carrying bi-directional promoters coordinately expressed two transgenes in the vast majority of transduced cells clearly outperforming the bicistronic vectors. The efficient performance of the new bi-directional LVs in primary hematopoietic cells, assayed ex vivo and after transplantation, and in several tissues in vivo, after direct vector delivery or transgenesis was established. The invention overcomes a long-standing hurdle in the quest for improved gene-expression tools and are expected to advance the reach and safety of gene therapy.

It is therefore an object of the instant invention a bidirectional promoter for expression of at least two coding sequences in opposite direction in animal cells comprising 5' end to 3' end:

a) a first minimal promoter sequence derived from cytomegalovirus (CMV) or mouse mammary tumor virus (MMTV) genomes;

b) a full efficient promoter sequence derived from an animal gene;

the two promoter sequences driving a coordinate transcription of said coding sequences in the opposite orientation.

In the ambit of the instant invention a full efficient promoter sequence means a sequence driving an efficient transcription of primary transcript. Preferably It comprises an enhancer region and a minimal promoter sequence, either distinct or overlapping. More preferably the full efficient promoter sequence derives from the phosphoglycerate kinase or from the ubiquitin promoter.

It is an object of the invention a bidirectional expression cassette essentially comprising the bidirectional promoter as above disclosed, convenient insertion sites positioned downstream to each promoter, and polyadenylation sites positioned downstream to each insertion site.

Preferably the bidirectional expression cassette further comprises at least one post-transcriptional regulatory element positioned upstream to one or each polyadenylation site. More preferably the bidirectional expression cassette further comprises at least one internal ribosome entry site (IRES) sequence to express three or more genes.

It is an object of the invention an expression construct containing the bidirectional promoter, as above disclosed.

It is an object of the invention an expression construct containing the bidirectional expression cassette, as above disclosed.

It is an object of the invention a gene transfer expression vector containing the expression construct as above disclosed further comprising lentiviral or retroviral sequences.

It is an object of the invention the use of the gene transfer expression vector for the delivery and expression of multiple genes in animal cells, preferably in vivo tissue animal cells, more preferably, brain neurons.

It is an object of the invention a method for the coordinate expression of two exogeneous coding sequences into an animal cell comprising the following steps:

a) cloning said coding sequences into the gene transfer expression vector according to claim 8, each coding sequence under the control of one of the two promoters of the bidirectional promoter;

b) transforming animal cells by means of said vectors;

c) allowing the expression of the vector.

Preferably the animal cell is a human cell, more preferably the human cell is a retransplantable human cell, even more preferably the retransplantable human cell is an hematopoietic cell.

Alternatively, the transformation of tissue cells in vivo may be performed by direct delivery of the vector, such as into brain neurons.

It is an object of the invention a method for generating a transgenic non human organism comprising the step of transforming appropriate cells by means of the gene transfer expression vector as disclosed above.

The vectors of the invention can be advantageously utilized for gene function and target validation studies in vitro and in vivo; gene therapy; expression of multiple genes in animal cells; generation of transgenic animals and eventually knock down of multiple genes; and for manufacturing of medicaments, as well.

FIGURE LEGENDS

The invention will be now described with reference to following Figures:

Fig. 1. Gene transfer performance of bicistronic lentiviral vectors. **(a)** Scheme of the proviral vector form. A bicistronic expression cassette containing an internal ribosome entry site (IRES) derived either from the encephalomyocarditis virus (EMCV), with wild-type (wt) or mutated (mut) translation start site, or from the 5' untranslated NF-kB repressing factor mRNA (NRF) was driven by the human immediate early cytomegalovirus (CMV) or phosphoglycerate kinase (PGK) promoter. ΔU3, R and U5, LTR regions with deletion in U3; SD and SA, splice donor and acceptor site; Ψ, encapsidation signal including the 5' portion of the gag gene (GA); RRE, Rev-response element; cPPT, central polypurine tract; WPRE, woodchuck hepatitis virus post-transcription regulatory element. **(b)** Southern blot analysis of HeLa cells transduced by the indicated monocistronic (CMV) or bicistronic vectors expressing luciferase (gene 1) and GFP (gene 2) from the CMV promoter, probed for the WPRE sequence. All vector integrated with the expected length of DNA. Vector copy number was determined relative to a plasmid standard curve and used to normalize vector stocks and ensure similar levels of integration for each vector in a given target cell type in the experiments shown in c-f. **(c-f)** Luciferase and GFP expression in human HeLa cells (c), umbilical vein endothelial cells (HUVEC, d), peripheral blood lymphocytes (PBL, e), and cord blood-derived CD34+ progenitors (f) transduced 5-7 days before with a monocistronic (□, CMV) or the indicated bicistronic CMV-luciferase-GFP vector. Left column, histograms representing net luciferase activity in cells extracts, mean ± SD. Right panel, dot plots representing GFP expression by FACS analysis, the frequency and the mean fluorescence intensity (MFI, X) of GFP+ cells is indicated. The control monocistronic vector expressed luciferase in the histogram (□), and GFP in the left-most dot plot (CMV) for each cell type. **(g, h)** FACS analysis of ΔNGFR and GFP expression in 293T cells (g) and CD34+ progenitors (h) transduced by a EMCV wt IRES vector expressing ΔNGFR and GFP from the PGK promoter. Histograms in panel (h) show the distribution of ΔNGFR expression in all viable cells analysed

(left), and of GFP expression in the gated (M1) ΔNGFR+ cells (right). Experiments shown are representative of at least three performed with similar results.

Fig. 2. Gene transfer performance of bidirectional lentiviral vectors. **(a)** Scheme of the proviral vector form. A bidirectional promoter made by minimal core promoter elements from the human cytomegalovirus (mCMV) joined upstream, and in opposite orientation, to an efficient promoter, derived from the human phosphoglycerate kinase (PGK) or poly-ubiquitin UBI-C gene, was driving divergent transcription of two RNAs. CTE, constitutive transport element from the Mason-Pfizer monkey virus; pA, polyadenylation site A from the Simian Virus 40. Other vector features as in the legend to figure 1. **(b)** Net luciferase activity and **(c-e)** GFP expression in HeLa cells transduced 5-7 days before with LVs carrying the indicated bi-directional or control expression cassettes. The frequency and MFI (X) of GFP+ cells at FACS analysis is indicated in the dot plots to the right. Luciferase activity was determined for the two marked vectors (□, ■). **(f-j)** ΔNGFR and GFP expression in HeLa cells transduced 5-7 days before with serial 10-fold dilutions of LVs carrying the indicated expression cassette. The frequency of ΔNGFR+ (upper left region) and ΔNGFR/GFP double positive (upper right region) cells, with the respective MFI of ΔNGFR (Y) and GFP (X), are indicated in the FACS dot plots. Experiments shown are representative of at least three performed with similar results.

Fig. 3. Comparison of bi-directional and bicistronic lentiviral vectors performance. ΔNGFR and GFP expression in 293T cells transduced 3 weeks before with serial 10-fold dilutions of LVs carrying the indicated expression cassette. The total percentage of ΔLNGFR-expressing cells and of ΔLNGFR/GFP double positive cells (in brackets) are indicated above the FACS dot plots. The average number of vector Copies per Cell (CpC) is indicated in each plot, with the expected frequency of transduced cells according to the Poisson's distribution of random independent events. Although virtually all integrated vectors expressed ΔNGFR, its level of expression and the fraction of transduced cells co-expressing GFP were much higher for the two bi-directional vectors tested (MA1 and MA4) as compared to the EMCV wt IRES bicistronic vector.

Fig. 4. Dual-gene transfer in hematopoietic cells by bi-directional vectors. **(a-c)** Human cord blood CD34+ progenitors were transduced by the GFP-ΔNGFR MA1 vector in the presence of early acting cytokines as described[23], and analysed either after 7 days of culture in the same medium (a), and after additional 10 days in medium promoting myeloid differentiation (b), or after seeding in methylcellulose-based clonogenic medium. For (a) and (b), a dot plot showing ΔNGFR and GFP expression by FACS analysis is shown, together with histograms showing the distribution of ΔNGFR expression in all viable cells analysed (top), and of GFP expression in the gated (M1) ΔNGFR+ cells (bottom). The percentage of immature progenitors expressing CD34, and of differentiating cells expressing the CD13 myeloid marker at the time of analysis is indicated. For (c), representative light (left) and fluorescent (right) micrograph of the indicated type of CFC are shown. **(d, e)** Human peripheral blood lymphocytes were transduced either after 2-day activation with anti-CD3 and anti-CD28 antibodies (d), or after 4-day treatment with interleukin-7, as described[24], (e), and analyzed for ΔNGFR and GFP expression as described above. **(f, g)** Purified (lin-) murine bone marrow progenitors were transduced without cytokine stimulation as described[48], and analyzed for ΔNGFR and GFP expression after 7 days in liquid culture (f), or immediately transplanted into lethally-irradiated syngenic recipients. FACS analysis of the peripheral blood of a representative mouse 2 months after transplant is shown in g. Experiments shown are representative of three performed with similar results. In d-f, cells transduced to low vector copy numbers are shown for more stringent performance analysis.

Fig. 5 In vivo dual-gene transfer by bi-directional vectors. High-titer of GFP-ΔNGFR MA1 LV were stereotactically injected into the striatum of adult mice. Cryostatic brain sections were obtained two months after injection and analyzed by immunofluorescence and confocal microscopy. Representative pictures of the injected area are shown, after immunostaining for ΔNGFR (red), GFP (green), and TO-PRO3 staining for nuclear DNA (blue). Fluorescent signals were sequentially acquired from single optical sections and are shown individually and after merging (merge). Original magnification 200X (Scale bar = 120 μm)

Fig. 6 Dual-transgenesis by bi-directional vector. Transgenic mouse lines were generated by direct injection of GFP-ΔNGFR MA1 LV into the perivitelline space of single-cell embryos, as described[19], and the indicated tissues were analyzed for ΔNGFR (red) and GFP (green) expression by immunofluorescence and confocal microscopy on cryostatic sections. Nuclei were stained by TO-PRO3 (blue). Fluorescent signals were sequentially acquired from single optical sections and are shown individually and after merging (merge). The pictures shown were obtained from an F1 mouse carrying two vector genomes integrated into the germ-line. Similar pictures were obtained from other transgenic mice analyzed that carried similar or higher number of vector copies. Original magnification 200X (spleen, lung), 400X (hearth, kidney, brain, liver), 630X (gut) (Scale bar = 120 μm)

Fig. 7a Map of the plasmid containing the lentiviral vector construct RRL-MA1-lucif/GFP.

Fig. 7b Sequence of the plasmid containing the lentiviral vector construct RRL-MA1-lucif/GFP.

Fig. 8a Map of the plasmid containing the lentiviral vector construct CCL-MA1-GFP/deltaLNGFR.

Fig. 8b Sequence of the plasmid containing the lentiviral vector construct CCL-MA1-GFP/deltaLNGFR.

Fig. 9a Map of the plasmid containing the lentiviral vector construct RRL-MA2-lucif/GFP.

Fig. 9b Sequence of the plasmid containing the lentiviral vector construct RRL-MA2-lucif/GFP.

Fig. 10a Map of the plasmid containing the lentiviral vector construct CCL-MA3-GFP/deltaLNGFR.

Fig. 10b Sequence of the plasmid containing the lentiviral vector construct CCL-MA3-GFP/deltaLNGFR.

Fig. 11a Map of the plasmid containing the lentiviral vector construct CCL-MA4-GFP/deltaLNGFR.

Fig. 11b Sequence of the plasmid containing the lentiviral vector construct CCL-MA4-GFP/deltaLNGFR.

EXAMPLE 1

MATERIALS AND METHODS

Plasmid construction

All transfer vectors were built from plasmid pCCL.sin.cPPT.PGK.GFP.WPRE [15] using the following previously described sequence elements: EMCV IRES's with the downstream gene coding sequence starting at the 11th ATG of the IRES (wt) or with the 11th ATG of IRES mutated to create a HindIII cloning site and allow translation initiation at the downstream transgene ATG [16](EMCVmut), the NRF IRES [18], the MPMV CTE[21], a minimal CMV core promoter[20], a 1226 bp fragment from the Ubiquitin-C promoter [19].

Construction of lentiviral vector with bi-directional promoters

To generate the lentiviral construct **RRL-MA1**, an XhoI-XhoI fragment containing the SV40polyA.CTE.Luciferase.minhCMV elements (derived from the lentiviral construct pRRL.sin.cPPT.SV40polyA.CTE.Luciferase.minhCMV.TetO7.minMMTV.eGFP) was cloned into the lentiviral vector construct pRRL.sin.cPPT.hPGK.eGFP.Wpre (Follenzi et al., 2000) cut with the same enzyme to obtain RRL-MA1-lucif/GFP (pRRL.sin.cPPT.SV40polyA.CTE.Luciferase.minhCMV.hPGK.eGFP.Wpre).

To generate the lentiviral construct **CCL-MA1**, two fragments were cloned into the lentiviral construct pRRL.sin.cPPT.hPGK.ΔLNGFRWpre first cut with KpnI, blunted and then cut with XhoI, the first fragment containing the minhCMV.eGFP elements was derived from the lentiviral construct pRRL.sin.cPPT.SV40polyA.CTE.Luciferase.minMMTV.TetO7.minhCMV.eGFP cut with KpnI, blunted and then with XhoI and the second derived from the construct pRRL.sin.cPPT.SV40polyA.CTE.tTA2.Wpre cut with BamHI, blunted and then cut with NotI. The resulting lentiviral construct pRRL.sin.cPPT.SV40polyA.CTE.Luciferase.minMMTV.TetO7.minhCMV.eGFP was cut with NotI and AvrII and the fragment containing the cPPT.SV40polyA.CTE.eGFP.minhCMV.hPGK. ΔLNGFRWpre was cloned into the lentiviral construct pCCL.sin.cPPT.hPGK.eGFP.Wpre cut with the same enzymes to

obtain CCL-MA1-GFP/ΔLNGFR (pCCL.sin.cPPT.SV40polyA.CTE.eGFP.minhCMV.hPGK. ΔLNGFRWpre).

To generate the lentiviral construct **RRL-MA2**, a HindIII-BamHI fragment containing the hPGK.Luciferase elements (derived from the lentiviral vector construct pRRL.sin.cPPT.hPGK.Luciferase.IRES.Wpre) was cloned into the retroviral construct SF2-cLCM2G (obtained from Rainer Loew, University of Heidelberg, FRG) cut with the same enzymes to obtain the construct cPPT.SV40polyA.CTE.Luciferase.hPGK.minMMTV.eGFP. This construct was first cut with SalI, blunted and then cut with BamHI and the fragment containing the Luciferase.hPGK.minMMTV.eGFP elements was cloned into the lentiviral vector construct pRRL.sin.cPPT.SV40polyA.CTE.tTA2.Wpre cut in the same way, to obtain RRL-MA2-lucif/GFP (pRRL.sin.cPPT.SV40polyA.CTE.Luciferase.hPGK.minMMTV.eGFP.Wpre).

To generate the lentiviral construct **CCL-MA3**, two fragments were cloned into the pBLKS+ cut with HindIII and XhoI, the first fragment containing the CTE.SV40polyA elements was derived from the lentiviral vector construct pRRL.sin.cPPT.SV40polyA.CTE.tTA2 cut with HindIII and XbaI and the second fragment containing the minMMTV.GFP elements derived from the construct cPPT.SV40polyA.CTE.Luciferase.hPGK.minMMTV.eGFP cut with XhoI and XbaI to obtain the construct pBLKS+ minMMTV.GFP.CTE.SV40polyA. The resulting construct was cut with EcoRV and XhoI and the fragment containing the minMMTV.GFP.CTE.SV40polyA was cloned into the lentiviral vector construct pCCL.sin.cPPT.hPGK. ΔNGFR.Wpre cut with the same enzymes,to obtain the final lentiviral vector construct CCL-MA3-GFP/ΔNGFR (pCCL.sin.cPPT.SV40polyA.CTE.GFP.minMMTV.hPGK. ΔNGFR.Wpre)

To generate the lentiviral construct **CCL-MA4** the fragment derived from pHR'.UBI-C.eGFP cut with PacI, blunted and cut with PstI, containing the UBI-C promoter sequence, was inserted into the place of the PGK promoter into construct pCCL.sin.cPPT.SV40polyA.CTE.GFP.minCMV.PGK.ΔNGFR.Wpre cut with EcoRV and PstI to obtain the final lentiviral vector construct CCL-MA4-

GFP/ΔNGFR (pCCL.sin.cPPT.SV40polyA.CTE.GFP.minCMV.UBI-C. ΔNGFR.Wpre)

The maps and the nucleotide sequences of the RRL-MA1-lucif/GFP, CCL-MA1-GFP/ΔLNGFR, RRL-MA2-lucif/GFP; CCL-MA3-GFP/ΔLNGFR; CCL-MA4-GFP/ΔLNGFR constructs are shown respectively in figures 7a-11a and figures 7b-11b.

Vector production and titration

VSV-pseudotyped third-generation LV were produced by transient 4-plasmid co-transfection into 293T cells and purified by ultracentrifugation as described [15], with the modification that 1 mM NaButyrate was added to the cultures for vector collection[47]. Expression titer of GFP or ΔLNGFR vectors were estimated on HeLa cells by limiting dilution. Vector particle was measured by HIV-1 gag p24 antigen immunocapture (NEN Life Science Products). Vector infectivity was calculated as the ratio between titer and particle for the vector expressing GFP or ΔNGFR. Vector expression titer in the 293T supernatant ranged from 0.7 to $1x10^7$ Transducing $Units^{HeLa}$(TU)/ml for monocistronic CMV or PGK vector, from 3 to $8x10^6$ TU/ml for bicistronic vectors and bi-directional vectors. Vector infectivity ranged from 0.5 to $1x10^5$ TU/ng of p24 for monocistronic CMV or PGK vector, and from 2 to $6x10^4$ TU/ng of p24 for bicistronic and bi-directional vectors.

Cell cultures

Continuous cultures of HeLa and 293T cells were maintained in Iscove's modified Dulbecco's medium (IMDM; Sigma, Milan, Italy) supplemented with 10% fetal bovine serum (FBS; Gibco, Invitrogen Corporation, UK) and a combination of penicillin-streptomycin and glutamine. Primary cultures of human umbilical vein endothelial cells (HUVECs), peripheral blood lymphocytes, and cord blood CD34+ progenitors were obtained and maintained as described[15]. $CD34^+$ progenitors were transduced with $5x10^7$ TU/ml of LV and cultured for at least 7 days in the presence of recombinant human interleukin 6 (rhIL6, 20 ng/ml), recombinant human stem cell factor (rhSCF, 100 ng/ml), recombinant human FLT-3 ligand (rhFLT-3 ligand, 100ng/ml), all from PeproTech (Rocky Hill, NJ), and recombinant human thrombopoietin (rhTPO, 20 ng/ml; Amgen, Thousand Oaks, CA) as described[23]. For

differentiating conditions, transduced progenitors were cultured for 10 days in the presence of rhSCF, 50 ng/ml, recombinant human granulocyte monocyte-colony stimulating factor (rhGM-CSF, 20 ng/ml), recombinant human monocyte-colony stimulating factor (rhG-CSF, 20 ng/ml), all from PeproTech. For clonogenic assays, transduced cells were plated at a density of 800 cells/ml in human complete MethoCult medium (StemCell Technologies, Vancouver, CA) and were scored by light and fluorescence microscopy 14 days later.

Human peripheral blood lymphocytes were purified by Ficoll gradient and transduced with 0.5-5x10^7 TU/ml of vector either after 2-day activation with 30ng/ml anti-CD3 antibodies (Orthoclone, Milan, Italy) plus 1μg/ml anti-CD28 antibodies (PharMingen, San Diego, CA), or after 4-day treatment with 5ng/ml interleukin-7 (Boehringer Mannheim-Roche GmbH, Mannheim, Germany), as described[24].

Purification of lineage marker-negative cells from C57BL/6 mouse bone marrow with a magnetic cell depletion technique (StemCell Technologies, Vancouver, CA), *ex vivo* transduction in serum-free StemSpan medium (StemCell Technologies, Vancouver, CA) with 0.5-2x10^7 TU/ml of vector, and transplantation into lethally irradiated syngenic recipients were performed as described [48].

Mice

CD1, C57BL/6 and FVB mice were purchased from Charles Rivers Laboratories (Calco, Italy) and maintained in SPF conditions. All animal procedures were performed according to protocols approved by the Hospital San Raffaele Institutional Animal Care and Use Committee.

DNA analysis: Southern and real time PCR

Vector copies per genome were quantified by Real-Time PCR from 300 ng template DNA extracted from cells by a commercial kit (Qiagen), using one set of primers and probe to detect the LV backbone:

LV forward primer, 5'-TGAAAGCGAAAGGGAAACCA-3';

LV reverse primer, 5'-CCGTGCGCGCTTCAG-3';

LV probe, 5'-(VIC)-CTCTCTCGACGCAGGACT-(TAMRA)-3'.

Reactions were carried out according to manufacturer instructions and analysed using the ABI Prism 7700 sequence detection system (PE-Applied Biosystem). For

Southern blot, DNA was extracted from transduced cells, digested with Afl-II to release the expression cassette from integrated vector DNA and analysed with a WPRE probe to detect vector sequences. The average number of integrated vector copies was determined relative to a plasmid standard curve.

These numbers were used to calculate vector integration titer and normalize vector stocks for all subsequent transduction experiments to ensure similar levels of integration for each vector tested.

Experimental Design and Stereotactic Injection.

Nine weeks-old C57BL/6 mice were anesthetized with intraperitoneal injection of Tribromoethanol 1.25% (SIGMA), positioned in a stereotactic frame (David Kopf Instruments, Tujunga, CA) and the skull exposed by a small incision. Two μl of vector concentrate (2 x10^6 TU/μl) was injected by a Hamilton syringe with a 33G blunt tip needle (Hamilton, Reno, NV) into the left hemisphere striatum (stereotactic coordinates in mm from bregma: AP=+0.74, ML=-1.9 and DV=-3.5 from skull surface) at a rate of 0.2 μl/min. The needle was left in place for additional 5 minutes before slow removal.

Transgenesis

Transgenic mice were generated using LV as described by Lois et al.[19]. Briefly, female FVB mice were superovulated with a combination of pregnant mare serum and human chorionic gonadotropin. On average between 20 and 30 embryos were collected per female and microinjected into the perivitelline space with 10-100 pL of 5x10^7 TU/ml LV stock on the same day. Manipulated embryos were immediately implanted into the oviduct of pseudopregnant CD1 mice. Pups were genotyped for the presence of the GFP sequence by PCR analysis as described[49]. Positive mice were bred to test germ-line trasmission of the transgene. DNA was extracted from the tail and used to quantify vector copy number by real time PCR in founder and F1 progeny mice.

Flow cytometry and Luciferase assay

Transduced cells were grown for at least 4 days before FACS analysis to reach steady state GFP expression and to rule out pseudotransduction. Before FACS analysis, adherent cells were detached with 0.05% trypsin-EDTA, washed, and fixed

in phosphate buffer saline (PBS) containing 1% paraformaldehyde (PAF) and 2% FBS. Cells grown in suspension were washed and resuspended in PBS containing 2 μg/ml propidium iodide (PI) (BD Bioscience PharMingen, San Diego, CA) and 2% FBS. For immunostaining, 10^5 cells were blocked in PBS 5% mouse serum, 5% human serum, 2% FBS for 15 min at 4°C. After blocking, 10μl of R-phycoerythrin (RPE)-conjugated antibodies (anti-CD34 and anti-CD13, Dako, Glostrup, Denmark, and anti-ΔLNGFR, BD Bioscience PharMingen, San Diego, CA) were added and the cells were incubated for 30 min at 4°C, washed, stained with PI, and analyzed by three-color flow cytometry. Only viable, PI-negative cells were used for the analysis. Luciferase was assayed in cell lysates prepared as described by the manufacturer (luciferase assay system, Promega). RLU were measured with a Lumat LB9507 luminometer (Berthold) after mixing cell lysates (normalized for protein content measured by BCA Protein Assay Reagent kit Pierce) with Luciferase Substrate (Promega).

Tissue analysis

Anesthetized mice were perfused with 0.9% NaCl followed by 4% PAF in PBS. Tissue samples were collected, equilibrated in 20% sucrose in PBS for 48 h at 4°C, and embedded in optimal-cutting-temperature compound (OCT) for quick freezing. 10μm (for transgenic mice) and 20μm (for stereotactic injected mice) thick cryostatic sections were post-fixed in PAF and frozen at –80 °C. Sections were blocked with 5% goat serum (Vector Laboratories) in PBS containing 1% bovine serum albumine (BSA) and 0.1% Triton X-100 (PBS-T), and incubated with rabbit affinity-purified GFP antibody (Molecular Probes) and R-phycoerythrin (RPE)-conjugated ΔLNGFR monoclonal antibody (BD Bioscience PharMingen, San Diego, CA) for 1 h, washed and stained with AlexaFluor488-conjugated goat anti-rabbit antibody (Molecular Probes) in PBS-T and 1% BSA for 1 h. Cell nuclei were stained with TOPRO-3 after 1h of RNAse treatment (Molecular Probes). Sections were mounted and analyzed by three-laser confocal microscope (Radiance 2100; BioRad). Fluorescent signals from single optical sections were sequentially acquired and analyzed by PhotoShop 7.0 (Adobe).

RESULTS

Bicistronic LVs

In order to express more than one transgene from a single vector, the authors first evaluated the performance of different IRES's in the context of late-generation self-inactivating LVs[15]. They used the strong CMV and PGK promoters to drive expression of bicistronic transcripts encoding, from the 5' to the 3' end, the luciferase reporter, an IRES, and the cell-associated GFP marker (fig. 1a). Two IRES's were derived from the Encephalomyocarditis virus; a wild-type (EMCVwt) and a mutant (EMCVmut) form [16, 17], that differed for the ATG from which downstream translation started. Another IRES was derived from the 5' untranslated sequence of the NF-kB transcription Repressing Factor (NRF) mRNA[18].

They generated high-titer VSV-pseudotyped stocks of all bicistronic and control monocistronic vectors, and normalized them for transducing activity measuring integration in HeLa cells by Southern blot (fig. 1b). They then compared gene expression in cells transduced to equal vector copy numbers (Fig. 1c-f). Although luciferase activity was similar in HeLa cells transduced by CMV-luciferase vector and in cells transduced by the best performing bicistronic vector, only a small fraction of the latter cells expressed the IRES-dependent GFP gene, with a ten-fold decrease in expression titer as compared to cells transduced by the control CMV-GFP vector (Fig. 1c). Moreover, the GFP mean fluorescence intensity (MFI) was significantly lower in cells expressing the protein from the IRES's than in cells expressing it from the ^{mRNA}Cap. They then tested bicistronic LVs in primary human cells, including umbilical vein endothelial cells, peripheral blood lymphocytes, and CD34+ cord blood hematopoietic progenitors (HPC) (Fig. 1d-f). All cell types were transduced efficiently, as indicated by the frequency of GFP-positive cells in cultures transduced by control CMV-GFP vector, but IRES-dependent GFP expression was only observed in a fraction of cells transduced by bicistronic vectors. IRES activity varied extensively with the target cell type; the NRF IRES was the only one reaching detectable downstream gene expression in lymphocytes, while the EMCVwt IRES was the most efficient in the other cell types. In addition, all IRES's decreased, in

some cases more than one log, upstream gene expression, as compared to the control CMV-luciferase vector.

They also evaluated IRES-based vectors by expressing two cell-associated markers, GFP and a truncated version of the low-affinity NGF receptor (ΔLNGFR) (Fig. 1g,h). Among HeLa cells transduced by a low dose of the best-performing bicistronic vector, only the cells expressing high levels of ΔLNGFR also expressed GFP, with an average of one out of four ΔNGFR-positive cells expressing GFP to detectable levels (Fig. 1g). Similarly, only a small fraction of transduced CD34+ progenitors expressing ΔNGFR also expressed GFP to detectable levels (Fig. 1h). Overall, these results indicated that IRES-based bicistronic vectors failed to ensure coordinate expression of two transgenes in most target cell types tested, and that multi-copy transduction or selection of transduced cells for downstream gene expression were required to obtain a population expressing both transgenes in the majority of cells.

Bidirectional LVs

To overcome the limitations of bicistronic vectors, the authors explored a new promoter design for coordinate transgene expression. They joined a minimal core promoter upstream, and in opposite orientation, to an efficient promoter. Rationale of this design was that upstream elements in the efficient promoter, when closely flanked by core promoters on both sides, may drive transcriptional activity in both directions. If such bi-directional activation occurred, expression of both transcripts would be coordinately regulated. They tested two ubiquitously expressed promoters, previously shown to drive robust and efficient transgene expression in LV; the above mentioned 516 bp fragment from the human phosphoglycerate kinase promoter (PGK) [15]and a 1226 bp fragment from the human ubiquitin C promoter (UBI C)[19]. They joined them to a minimal core promoter derived from the cytomegalovirus (minCMV) that was previously developed to couple initiation of eukaryotic transcription to tetracycline (Tc)-dependent operators[20]. They flanked the bi-directional promoter with two expression cassettes optimized for LV-mediated gene delivery (fig. 2a). The upstream cassette - in anti-sense orientation relative to the vector LTR – included the constitutive transport element (CTE) of the Mason-Pfizer virus[21], and a polyadenylation site from the Simian Virus 40 (SV40). The

downstream cassette included the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE)[22] and the SIN HIV-1 LTR polyadenylation site.

As described above for bicistronic LVs, they verified correct transfer and normalized transduction of each vector by Southern blot analysis and real-time PCR of transduced cells. LV carrying bi-directional expression cassettes were produced to high titer and infectivity, similar to those obtained with standard vectors (see Methods). The bi-directional design significantly enhanced transcription from the upstream minimal promoter without affecting downstream expression from the efficient promoter (fig. 2b-h). Luciferase expression from the minCMV promoter, for instance, was increased at least one log when fused upstream to the PGK promoter (fig. 2b). Remarkably, the bi-directional PGK promoter allowed detecting GFP (or ΔLNGFR, not shown) to the same frequency and to similar expression levels in cells transduced by the bi-directional vector and expressing the protein from either side of the promoter (fig. 2c,d), as in cells transduced by the control PGK vector (fig. 2e). Using two cell-associated markers, ΔLNGFR and GFP, they showed stable, efficient and coordinate expression of bi-directional LVs, both at high and low vector copy number (fig. 2f). At high vector input, they reached high-level expression of both transgenes in virtually every target cell. At low vector input, when most transduced cells carried one proviral copy, they showed transgene co-expression in virtually every labeled cell, indicating the occurrence of divergent transcription from the bi-directional promoter. In both conditions, transgene expression was maintained to similar levels in cells analyzed at early and late times post-transduction (not shown, and Fig. 3 below). Transgene-expressing cells tended to distribute along a diagonal line in the two-color FACS plot, indicating that expression of the two transgenes was coordinately regulated.

Intriguingly, they observed coordinate bi-directional expression, although to significantly lower efficiency on the upstream side than the downstream side, when they tested the sole PGK promoter in the context of the bi-directional expression cassette that they developed (fig. 2g). They reproduced this finding after swapping the position of the two transgenes on each sides of the PGK promoter (not shown). These results indicated that transcription-activating elements in the PGK promoter

are intrinsically capable of triggering divergent transcription and thus provide the main driving force for dual-gene expression in the new LV, ensuring coordinate regulation of transcription on both sides of the bi-directional promoter. Apposition of the minCMV core promoter, which had a very low activity per se (fig. 2h, and 2b above), enhanced upstream transcription from the PGK promoter possibly because of more efficient initiation (compare fig. 2g and 2f). When they changed the driving promoter in bi-directional vectors from PGK to UBI-C, they reproduced the findings observed with the PGK promoter (fig. 2i). They revealed an intrinsic bi-directional activity of the UBI-C promoter (fig. 2j) that was significantly enhanced by the upstream addition of the minCMV promoter.

They then compared directly the performance of bi-directional and bicistronic vectors in relation to the number of integrated copies, as measured by real-time PCR (Fig. 3). By analyzing 293T cells transduced with increasing vector doses, they proved that the vast majority of integrated bi-directional vectors based on the PGK (MA1) or UBI-C (MA4) promoter efficiently expressed both transgenes, clearly outperforming the best IRES-based bicistronic vector.

Ex Vivo and In Vivo Dual-Gene Transfer

They then assessed the performance of the bi-directional MA1 LV in more relevant targets for gene therapy applications and by different delivery strategies. They transduced human cord-blood HPC and PBL with ΔLNGFR-GFP MA1 LV *ex vivo*, according to previously optimized protocols[23, 24] (Fig.4). Both gene products were coordinately expressed to high-levels in a large fraction of HPC scored both as immature cells grown in the presence of early-acting cytokines (Fig. 4a), and after differentiation in liquid culture (Fig. 4b) or clonogenic assay (Fig. 4c, GFP only). Similarly, they obtained coordinate ΔLNGFR and GFP expression in PBL transduced in standard conditions of proliferation, triggered by CD3/CD28 co-stimulation (Fig. 4d), and as non-proliferating cells, treated only with IL-7 to maintain naïve cell properties (Fig. 4e). They also performed transplantation studies with transduced murine HPC, enriched from the bone marrow by negative selection, to prove stable dual-transgene expression in the progeny of long-term repopulating HSC (Fig 4f). ΔLNGFR and GFP were coordinately expressed to similar levels in the

ex vivo transduced cells, before transplantation, and in the white blood cells of long-term engrafted mice. Overall, these results validated the new LV for proficient dual gene transfer in primitive, committed, and differentiated hematopoietic cells.

They injected concentrated ΔLNGFR-GFP MA1 LV in the striatum of adult mice and scored transgene expression 4 weeks after injection by confocal microscopy of brain sections immuno-stained for GFP and ΔLNGFR (fig. 5). They observed robust co-expression of both transgenes in the brain tissue surrounding the injection site. As previously reported after striatal injection of VSV-pseudotyped LV[25-27], the vast majority of cells expressing the markers had the typical morphology of striatal neurons. Thus, the new bi-directional LV enabled efficient *in vivo* dual-gene transfer.

Dual-Transgenesis

They evaluated whether the new bi-directional LV allowed generation of dual-transgenic mouse lines. As previously described by Lois et al[19], they microinjected the ΔNGFR-GFP LV into the perivitelline space of single-cell embryos, and implanted them into pseudopregnant females. We obtained transgenic mice to high frequency, as assessed by the presence of vector DNA (more than 50% of newborns), and proved vector integration in the germ line by crossing some founder mice and analyzing their progeny for vector DNA content and transgene expression (Fig. 6). In the two F1 mice analyzed, carrying 2 and 5 vector copies in the genome, they found remarkably consistent expression of both transgenes in virtually every cell in the tissues studied, which included brain, liver, spleen, gut, heart, skeletal muscle, and kidney. Vector expression was also well detectable in the bone marrow and peripheral blood of the same mice, although in less than 100% of the cells, and more clearly for ΔNGFR than GFP (not shown). These data indicated that bi-directional LV transgenesis is a rapid and efficient method to obtain robust, stable and coordinate expression of two transgenes in genetically-engineered mice. In addition, they show that the minCMV-PGK bi-directional promoter that they developed governs dual transgene expression in the majority of differentiated tissues of the mouse, and maintains expression after inheritance through the germ-line.

DISCUSSION

In the pursuit of strategies enabling efficient dual-gene transfer, they initially faced significant limitations of IRES-based approaches. When tested in the context of bicistronic LV, IRES-dependent gene expression was significantly lower than that dependent on the ^{mRNA}Cap, and required multi-copy transduction to co-express the downstream gene in a sizable fraction of transduced cells. In addition, IRES's decreased expression of the upstream gene in the transcript, and displayed significant cell type-dependent variation in activity. Similar limitations have been reported when incorporating IRES's into other types of gene transfer vectors[14, 28-32]. Thus, selection for downstream gene expression is likely to be required when using IRES to ensure co-expression in all target cells. Although selection protocols are compatible with some *ex vivo* gene transfer and therapy applications, they may adversely affect the biological properties of gene-corrected cells, in particular when selectable marker expression is inefficient. In fact, prolonged *ex vivo* culture and a limited size or clonal composition of the transduced cell population may reduce engraftment, long-term survival and tissue repopulation after transplantation[33]. Even more important, the inefficiency of IRES-dependent expression prevents most application of bicistronic vectors to direct *in vivo* gene transfer. Thus, authors explored novel strategies to take full advantage of gene transfer systems, such as LV, that allow efficient *ex vivo* transduction and direct *in vivo* administration[34].

They have developed a new promoter design based on the juxtaposition of core promoter elements upstream, and in opposite orientation, to an efficient promoter. The bi-directional assembly drove divergent transcription, indicating that upstream enhancer/promoter elements within the efficient promoter were capable of promoting transcription in an orientation-independent manner and from both sides simultaneously. Upon incorporation of these promoters into LV, they reached efficient dual-gene transfer and coordinate expression in continuous cell lines and primary cells *ex vivo*. Because both transgenes were expressed in the vast majority of transduced cells, they did not need to select cells to ensure transgene co-expression. Upon direct injection of bi-directional LV into the CNS, the authors showed coordinate expression of two transgenes in neural cells *in vivo*. In addition, bi-

directional LV allowed robust dual transgenesis, leading to pan-cellular expression of both transgenes in all tissues examined. All these results could not be reached until now using currently available technologies.

By monitoring transduced cells carrying a single vector copy, authors proved that divergent transcription occurred from a single bi-directional promoter, that expression of both transgenes was functionally linked and coordinately regulated, and that bi-directional promoters were consistently active in all types of target cells tested, without being silenced or randomly fixed in one direction of transcription, even after cellular differentiation. Although they did not map how close the two opposite core promoters must be for operational linkage, they may expect that close juxtaposition of the fused minimal core promoter to some of the upstream elements in the efficient promoter, as observed in natural promoters between core and upstream elements, may be required. Both the PGK and UBI-C promoters tested in this work drove divergent transcription when fused to a minimal core promoter in the opposite orientation. Intriguingly, both of these promoters were shown to be intrinsically capable of promoting divergent transcription, although to lower efficiency on the upstream than the downstream side, when incorporated into the bi-directional expression cassette that they developed. This surprising observation may indicate a specific feature of a class of ubiquitously-expressed housekeeping promoters, possibly related to their content of CpG islands (see below and[35-37]). However, they should not forget that both the promoter placement between two efficient expression cassettes endowed with post-transcriptional regulatory elements enhancing translation, and LV-mediated integration, which has been shown to preferentially target transcribed genes in the chromatin, may contribute to unravel latent transcriptional activity. Although the intrinsic bi-directional activity of the housekeeping promoters tested may not be efficient enough for exploitation *per se*, without the upstream assembly of core promoter elements described in this work, it provides the basis for the coordinate regulation of dual-gene expression reached by our new vectors. On the other hand, the propensity of these promoters to drive divergent transcription should be kept in mind when engineering vectors and analyzing transduced cells or tissues [38], and may provide a possible mechanism for

the frequently observed interference between nearby promoters in the same vector construct [10, 39]. It is possible that the bi-directional design described here may be successfully applied to tissue-specific promoters to obtain coordinated expression of two transgenes in specific tissues. In addition, by combining bi-directional promoters with bicistronic transcripts one could express more than two transgenes within the same cell, although with the limitations described above for IRES-dependent vectors. Inducible bi-directional promoters were originally developed in Tet-regulated expression systems, by duplicating a minimal promoter on both sides of a series of Tet operator repeats, to obtain exogenously regulated expression of two transgenes[36, 40, 41]. This design was recently applied to other systems that also combine prokaryotic enhancer elements with chimeric trans-activators to regulate gene expression[42]. Although these inducible expression systems represent powerful tools for gene-function studies, they are dependent on co-expression and functional activity of protein trans-activators, and pose several challenges when applied to vector-based delivery and *in vivo* applications. A constitutive bi-directional promoter was recently tested for exogenous gene expression in plant biotechnology[43]. Our results provide the first description of synthetic bi-directional promoters that exploit the endogenous transcriptional machinery available to most animal cell types to drive robust and constitutive expression of two divergent transcripts. In nature, few instances of bi-directional promoters had been documented until recently. Intriguingly, a recent survey of the human genome indicated an abundance of divergently transcribed gene pairs, whose transcription start sites are separated by less than 1 kb[44, 45]. It is likely that many of the promoter elements found between these gene pairs can initiate transcription in both directions, and contain shared elements that regulate both genes[46]. Thus, the synthetic bi-directional promoters that they have developed may mimic a well-represented and evolutionary conserved feature of eukaryotic transcription, providing a structural basis for their robust performance. The new lentiviral vectors built around these bi-directional promoters will likely advance the reach and the safety of gene therapy, the power of gene-function and target validation studies, and the applications of animal transgenesis. If adapted for the

expression of short interfering RNA, they may also enable coordinate knock-down of multiple genes.

REFERENCES

1. Kay, M.A., Glorioso, J.C. & Naldini, L. *Nat Med* 7, 33-40 (2001).
2. Neff, T. et al. *J Clin Invest* 112, 1581-1588 (2003).
3. Bordignon, C. & Roncarolo, M.G.. *Nat.Immunol.* 3, 318-321 (2002).
4. Sadelain, M.. *J Gene Med* 4, 113-121 (2002).
5. Bonini, C. et al. *Science* 276, 1719-1724 (1997).
6. Hacein-Bey-Abina, S. et al.. *Science* 302, 415-419 (2003).
7. Burton, E.A., Glorioso, J.C. & Fink, D.J *Gene Ther* 10, 1721-1727 (2003).
8. Sadelain, M., Riviere, I. & Brentjens, R.. *Nat Rev Cancer* 3, 35-45 (2003).
9. Miller, A.D. in Retroviruses. (eds. J. Coffin, S.H. Hughes & H.E. Varmus) 437-474 (Cold Spring Harbor Laboratory Press, Plainview; 2000).
10. Emerman, M. & Temin, H.M. *Mol Cell Biol* 6, 792-800 (1986).
11. Zhu, Y., et al.. *Mol.Ther.* 4, 375-382 (2001).
12. Klump, H. et al. *Gene Ther.* 8, 811-817 (2001).
13. Furler, S., et al.. *Gene Ther* 8, 864-873 (2001).
14. Martinez-Salas, E.. *Curr.Opin.Biotechnol.* 10, 458-464 (1999).
15. Follenzi, A., et al.. *Nat Genet* 25, 217-222 (2000).
16. Ghattas, I.R., Sanes, J.R. & Majors, J.E.. *Mol.Cell Biol.* 11, 5848-5859 (1991).
17. Qiao, J., et al.. *Hum Gene Ther* 13, 881-887 (2002).
18. Oumard, A., et al.. 20, 2755-2759 (2000).
19. Lois, C., et al.. *Science* 295, 868-872 (2002).
20. Baron, U. & Bujard, H.. *Methods Enzymol.* 327, 401-421 (2000).
21. Bray, M. et al. *Proc Natl Acad Sci U S A* 91, 1256-1260 (1994).
22. Zufferey, R., et al.. *J.Virol.* 73, 2886-2892 (1999).
23. Ailles, L. et al.. *Mol Ther* 6, 615-626 (2002).
24. Cavalieri, S. et al.. *Blood* 102, 497-505 (2003).
25. Naldini, L., et al. *Proc Natl Acad Sci U S A* 93, 11382-11388 (1996).
26. Baekelandt, V. et al. *Hum Gene Ther* 13, 841-853 (2002).

27. Deglon, N. et al.. *Hum.Gene Ther.* 11, 179-190 (2000).

28. Sokolic, R.A. et al. *Blood* 87, 42-50 (1996).

29. Wong, E.T., Ngoi, S.M. & Lee, C.G. *Gene Ther* 9, 337-344 (2002).

30. Kozak, M.. *Gene* 318, 1-23 (2003).

31. Mizuguchi, H., et al.. *Mol Ther* 1, 376-382 (2000).

32. Hennecke, M. et al.. *Nucleic Acids Res* 29, 3327-3334 (2001).

33. Mazurier, F., et al. *Blood* 103, 545-552 (2004).

34. Vigna, E. & Naldini, L.. *J Gene Med* 2, 308-316 (2000).

35. Gardiner-Garden, M. & Frommer, M.. *J Mol Biol* 196, 261-282 (1987).

36. Lavia, P., Macleod, D. & Bird, A.. *EMBO J* 6, 2773-2779 (1987).

37. Johnson, P. & Friedmann, T.. *Gene* 88, 207-213 (1990).

38. Scacheri, P.C. et al. *Genesis* 30, 259-263 (2001).

39. Vigna, E. et al.. *Mol Ther* 5, 252-261 (2002).

40. Baron, U., et al.. *Nucleic.Acids.Res.* 23, 3605-3606 (1995).

41. Unsinger, J., et al.. *Mol Ther* 4, 484-489 (2001).

42. Fux, C. et al.. *J Gene Med* 5, 1067-1079 (2003).

43. Xie, M., He, Y. & Gan, S*Nat Biotechnol* 19, 677-679 (2001).

44. Trinklein, N.D. et al.. *Genome Res* 14, 62-66 (2004).

45. Takai, D. & Jones, P.A. *Mol Biol Evol* 21, 463-467 (2004).

46. Adachi, N. & Lieber, M.R.. *Cell* 109, 807-809 (2002).

47. Farson, D. et al. *Hum Gene Ther* 12, 981-997 (2001).

48. De Palma, M., et al.. *Nat Med* 9, 789-795 (2003).

49. Follenzi, A., et al. *Hum Gene Ther* 13, 243-260 (2002).